7-1-2022

Monotonic Bit-Invariant Permutation-Invariant Metrics on the Set of All Infinite Binary Sequences

Irina Padilla
University of Ostrava, Irina.Perfilieva@osu.cz

Vladik Kreinovich
The University of Texas at El Paso, vladik@utep.edu

Comments:
Technical Report: UTEP-CS-22-86

Monotonic Bit-Invariant Permutation-Invariant Metrics on the Set of All Infinite Binary Sequences

Irina Perfilieva and Vladik Kreinovich

Abstract In a computer, all the information about an object is described by a sequence of 0s and 1s. At any given moment of time, we only have partial information, but as we perform more measurements and observations, we get longer and longer sequence that provides a more and more accurate description of the object. In the limit, we get a perfect description by an infinite binary sequence. If the objects are similar, measurement results are similar, so the resulting binary sequences are similar. Thus, to gauge similarity of two objects, a reasonable idea is to define an appropriate metric on the set of all infinite binary sequences. Several such metrics have been proposed, but their limitation is that while the order of the bits is rather irrelevant – if we have several simultaneous measurements, we can place them in the computer in different order – the distance measured by current formulas change if we select a different order. It is therefore natural to look for permutation-invariant metrics, i.e., distances that do not change if we select different orders. In this paper, we provide a full description of all such metrics. We also explain thee limitation of these new metrics: that they are, in general, not computable.

1 Formulation of the Problem

Need for a metric. To gain knowledge about an object or a system, we perform measurements and observations. Nowadays, the results of both measurements and observations are stored in a computer, and in a computer, everything is represented as a binary sequence, i.e., as sequence of 0s and 1s. Thus, our information about

Irina Perfilieva
Institute for Research and Applications of Fuzzy Modeling, University of Ostrava, 30. dubna 22, 701 03 Ostrava 1, Czech Republic, Irina.Perfilieva@osu.cz

Vladik Kreinovich
Department of Computer Science, University of Texas at El Paso, 500 W. University
El Paso, Texas 79968, USA, e-mail: vladik@utep.edu

an object can be represented as a potentially infinite binary sequence – potentially infinite, since, to gain more information, we can always perform more measurements.

A natural way to describe closeness between the two objects is by comparing the corresponding binary sequences: if two objects are similar, the results should be similar, so the corresponding sequences should be close. To describe this similarity, it is therefore reasonable to have a metric on the set $B = \{0,1\}^{N^+}$ of all infinite binary sequences $x = x_1 x_2 \ldots x_n \ldots$

Comment on notations. Here, we denoted:

- by $N^+ = \{1, 2, \ldots\}$, the set of all positive integers, and
- by A^B we – as usual – denote the set of all the functions from the set B to the set A.

What is a metric: reminder. What are the reasonable properties of this metric? To formulate these properties, let us first recall the usual definition of a metric.

Definition 1. *Let X be a set. A* metric *is a mapping $d : X \times X \to \mathbb{R}_0^+$ that assigns, to every pair (x, y) of elements of the set X, a non-negative number $d(x, y)$, and that satisfies the following three properties for all elements x, y, and z:*

- $d(x, y) = 0$ *if and only if* $x = y$;
- $d(x, y) = d(y, x)$ (symmetry), *and*
- $d(x, z) \le d(x, y) + d(y, z)$ (triangle inequality).

First reasonable property: bit invariance. When we represent observation results in a computer, what is important is that for each bit, we have two different values. It is not important which of these values is associated with 1 and which with 0. For example, to describe whether a person is sick, we have the following two options:

- We can ask whether the person is sick. In this case, a sick person corresponds to the "yes" answer, i.e., to 1.
- Alternatively, we can ask whether the person is healthy. In this case, a sick person corresponds to the "no" answer, i.e., to 0.

Since selecting 0 or 1 for each n is arbitrary, it is reasonable to require that the metric does not change if we simply swap some 0s and 1s. Let us describe this property in precise terms.

Definition 2.

- *Let $S \subseteq N^+$ be a subset of the set of all positive integers. For each infinite binary sequence x, by its S-*swap *$B_S(x)$, we mean a sequence $y = y_1 y_2 \ldots$ where:*
  - *for $n \in S$, we have $y_n = 1 - x_n$, and*
  - *for $n \notin S$, we have $y_n = x_n$.*
- *We say that a metric $d : B \times B \to \mathbb{R}_0^+$ is* bit-invariant *if for every set $S \subseteq N^+$ and for every two sequences $x, y \in B$, we have*

$$d(x, y) = d(B_S(x), B_S(y)).$$

Comment on notations. By $\mathbb{R}_0^+$, we denote the set of all non-negative real numbers.

Second reasonable property: monotonicity. Suppose that we have three sequences x, y, and z such that z differs from x in all places in which y differs from x and maybe in other places as well. In other words, z is "more different" from x than y. In this case, it is reasonable to require that the distance between x and z is larger – or at least the same, but not smaller – than the distance between x and y.

Let us describe this reasonable property in precise terms.

Definition 3. *We say that a metric $d : B \times B \to \mathbb{R}_0^+$ is* monotonic *if for all triples of sequences (x, y, z) for which $x_n \neq y_n$ implies $x_n \neq z_n$, we have $d(x, y) \leq d(x, z)$.*

Examples of metrics that satisfy both properties. There are many metrics on the set B of all infinite binary sequences that are both bit-invariant and monotonic. The most well-known example (see, e.g., [1, 2]) is the metric

$$d(x, y) = \sum_{n=1}^{\infty} 2^{-n} \cdot |x_n - y_n|. \tag{1}$$

It is also possible to have a more general family of bit-invariant and monotonic metrics

$$d(x, y) = \sum_{n=1}^{\infty} q^n \cdot |x_n - y_n| \tag{2}$$

corresponding to different values $q \in (0, 1)$.

Limitations of the known metrics. The problem with these metrics is related to the fact that the numbering of the bits in a binary sequence is rather arbitrary: if several sensors measured some quantities at the same time, we can place the results of the 1st sensor first or the result of the 2nd sensor first, etc. From this viewpoint, the closeness $d(x, y)$ of the two objects should not depend on which order we choose, i.e., should not change if we apply the same permutation to the sequences x and y.

However, the above two metrics – and most other metrics – change if we permute the bits. For example, in the expression (1), the difference $|x_1 - y_1|$ between the first bits enters with the coefficient $1/2$, while the difference $|x_2 - y_2|$ between the second bits enters with the coefficient 1/4. So, if we swap the first and the second bits, we get, in general, a different distance value.

Resulting problem and what we do in this paper. Because of the above limitation of known metrics, it is desirable to describe permutation-invariant metrics. This is what we do in this paper.

2 Definitions and the Main Result

Definition 4.

- *By a* permutation*, we mean a 1-1 mapping $\pi : N^+ \to N^+$ for which $\pi(N^+) = N^+$.*

- *By the* result $\pi(x)$ *of applying a permutation* π *to an infinite sequence* $x = x_1x_2\ldots$, *we means a sequence* $x_{\pi(1)}x_{\pi(2)}\ldots$

Definition 5. *We say that a metric* $d : B \times B \to \mathbb{R}_0^+$ *is* permutation-invariant *if for every two sequences* x *and* y *and for every permutation* π*, we have*

$$d(\pi(x), \pi(y)) = d(x, y).$$

Definition 6.

- *We say that the sequences* x *and* y have k different bits *if there are exactly* k *indices* n *for which* $x_n \neq y_n$.
- *We say that the sequences* x *and* y have k common bits *if there are exactly* k *indices* n *for which* $x_n = y_n$.

Proposition 1. *For a metric* d *on the set* B *of all infinite binary sequences, the following two conditions are equivalent to each other:*

- *the metric* d *is monotonic, bit-invariant, and permutation-invariant;*
- *there exists real numbers*

  $$0 = d_0^{\neq} \le d_1^{\neq} \le \ldots \le d_k^{\neq} \le \ldots \le d_\infty \le \ldots d_k^{=} \le \ldots \le d_2^{=} \le d_1^{=} \le d_0^{=}$$

  for which

  $$d_{k+\ell}^{\neq} \le d_k^{\neq} + d_\ell^{\neq} \text{ and } d_{k-\ell}^{=} \le d_k^{=} + d_\ell^{\neq} \tag{3}$$

  so that:

  – *when* x *and* y *have* k *different bits, then* $d(x, y) = d_k^{\neq}$;
  – *when* x *and* y *have* k *common bits, then* $d(x, y) = d_k^{=}$; *and*
  – *when* x *and* y *have infinitely many different bits and infinitely many common bits, then* $d(x, y) = d_\infty$.

Example. One can check that the desired inequalities are satisfied for $d_k^{\neq} = 1 - 2^{-k}$, $d_\infty = 1$, and $d_k^{=} = 1 + 2^{-k}$. Indeed:

- The first of the two inequalities (3) takes the form

  $$1 - 2^{-(k+\ell)} \le 1 - 2^{-k} + 1 - 2^{-\ell},$$

  i.e., if we denote $a \stackrel{\text{def}}{=} 2^{-k}$ and $b \stackrel{\text{def}}{=} 2^{-\ell}$, the form $1 - a \cdot b \le 1 - a + 1 - b$. If we subtract the left-hand side from both sides of this inequality, we get an equivalent inequality $0 \le 1 - a - b + a \cdot b$, i.e., $0 \le (1 - a) \cdot (1 - b)$, which is, of course, always true since $a = 2^{-k} \le 1$ and $b = 2^{-\ell} \le 1$.
- The second of the two inequalities (3) takes the form

  $$1 + 2^{-(k-\ell)} \le 1 + 2^{-k} + 1 - 2^{-\ell},$$

i.e., if we denote $a \stackrel{\text{def}}{=} 2^{-(k-\ell)}$ and $b \stackrel{\text{def}}{=} 2^{-\ell}$, the form $1+a \le 1+a\cdot b+1-b$. If we subtract the left-hand side from both sides of this inequality, we get an equivalent inequality $0 \le 1-a-b+a\cdot b$, i.e., $0 \le (1-a)\cdot(1-b)$, which is always true.

Proof.

1°. One can easily check that the function $d(x,y)$ described by the second condition is indeed a monotonic bit-invariant and permutation-invariant metric.

Specifically, monotonicity and invariances are easy to prove. The only thing that is not fully trivial to prove is that d as a metric satisfies the triangle inequality. Below, we propose the proof.

1.1°. If:

- x differs from y in k places – so that $d(x,y) = d^{\ne}_k$, and
- y differs from z in ℓ places – so that $d(y,z) = d^{\ne}_\ell$,

then x and z can differ in no more than $k+\ell$ places, so $d(x,z) = d^{\ne}_m$ for some $m \le k+\ell$. Due to the inequalities between the values $d^{\ne}_k$, we have $d(x,z) = d^{\ne}_m \le d^{\ne}_{k+\ell}$, and due to (3), we have

$$d(x,z) \le d^{\ne}_{k+\ell} \le d^{\ne}_k + d^{\ne}_\ell = d(x,y) + d(y,z),$$

exactly what we wanted to prove.

1.2°. If:

- x coincides with y in k places – so that $d(x,y) = d^{=}_k$, and
- y differs from z in ℓ places – so that $d(y,z) = d^{\ne}_\ell$,

then, if $k > \ell$, x and z coincide in at least $k-\ell$ places, so $d(x,z) = d^{=}_m$ for some $m \ge k-\ell$. Due to inequalities between the values $d^{=}_k$, we have $d(x,z) = d^{=}_m \le d^{\ne}_{k-\ell}$, and due to (3), we have

$$d(x,z) \le d^{=}_{k-\ell} \le d^{=}_k + d^{\ne}_\ell = d(x,y) + d(y,z),$$

exactly what we wanted to prove.

1.3°. For the cases:

- when at least one of the pairs (x,y) or (y,z) has infinitely many coinciding and infinitely many differing indices, and
- when both pairs coincide in finitely many places,

triangle inequality can be proven similarly.

2°. Let us now show that, vice versa, every monotonic bit-invariant and permutation-invariant metric on the set of all infinite binary sequences has this form. Specifically, we will show that the metric has this form for the following values:

$$d^{\ne}_k \stackrel{\text{def}}{=} d(00\ldots0\ldots, 1\ldots1\ (k \text{ times})\ 0\ldots0\ldots);$$

$$d_k^= \stackrel{\text{def}}{=} d(00\ldots0\ldots, 0\ldots0\ (k \text{ times})\ 1\ldots1\ldots); \text{ and}$$

$$d_\infty \stackrel{\text{def}}{=} d(00\ldots0\ldots, 0101\ldots01\ldots).$$

3°. Let us first use bit-invariance. In general, when we apply the same transformation T_S to two sequences x and y, resulting in $X = T_S(x)$ and $Y = T_S(y)$, then for each n:

- if $x_n = y_n$ then $X_n = Y_n$, and
- if $x_n \neq y_n$ then $X_n \neq Y_n$.

Let us take, as the set S, the set of all indices n for which $x_n = 1$. Then, the sequence $X = T_S(x)$ has the form $X = 00\ldots0\ldots$, and for each n, X_n and Y_n have equal or different values depending on whether x_n or y_n were equal or not. Due to bit-invariance, we have $d(x, y) = d(X, Y)$.

4°. Let us now use permutation-invariance, according to which

$$d(X, Y) = d(\pi(X), \pi(Y))$$

for all permutations π.

4.1°. If the sequences x and y differ in exactly k places, then X and Y also differ in exactly k places. Since the values X_n are all zeros, this means that the sequence Y is equal to 1 in exactly k places. We can thus perform a permutation π that moves these k places to the first k locations $1, \ldots, k$. This permutation does not change the all-zeros sequence X, so we have

$$d(x, y) = d(X, Y) = d(X, \pi(Y)) = d(00\ldots0\ldots, 1\ldots1\ (k \text{ times})\ 0\ldots0\ldots),$$

i.e., $d(x, y) = d_k^{\neq}$.

4.2°. If the sequences x and y coincide in exactly k places, then X and Y also coincide in exactly k places. Since the values X_n are all zeros, this means that the sequence Y is equal to 0 in exactly k places. We can thus perform a permutation π that moves these k places to the first k locations $1, \ldots, k$. This permutation does not change the all-zeros sequence X, so we have

$$d(x, y) = d(X, Y) = d(X, \pi(Y)) = d(00\ldots0\ldots, 0\ldots0\ (k \text{ times})\ 1\ldots1\ldots),$$

i.e., $d(x, y) = d_k^=$.

4.3°. Finally, if the sequences x and y differ in infinitely many places and coincide in infinitely many places, then X and Y also differ in infinitely many places and coincide in infinitely many places. Since the values X_n are all zeros, this means that the sequence Y has infinitely many 0s and infinitely many 1s. We can thus perform a permutation π that moves these all coinciding places into odd locations and all differing places into even locations. This permutation does not change the all-zeros sequence X, so we have

$$d(x,y)=d(X,Y)=d(X,\pi(Y))=d(00\ldots0\ldots,0101\ldots01\ldots),$$

i.e., $d(x,y)=d_\infty$.

5°. Let us now prove that the values $d_k^{\neq}$ and $d_k^{=}$ satisfy the desired inequalities (3).

5.1°. Let us first prove the first of the two inequalities (3).

Triangle inequality implies that

$$d(00\ldots0\ldots,1\ldots1\ (k+\ell\text{ times})\ 0\ldots0\ldots)\le$$
$$d(00\ldots0\ldots,1\ldots1\ (k\text{ times})\ 0\ldots0\ldots)+$$
$$d(1\ldots1\ (k\text{ times})\ 0\ldots0\ldots,1\ldots1\ (k+\ell\text{ times})\ 0\ldots0\ldots). \tag{4}$$

The first two distances in this equality are $d_{k+\ell}^{\neq}$ and $d_k^{\neq}$, so

$$d_{k+\ell}^{\neq}\le d_k^{\neq}+d(1\ldots1\ (k\text{ times})\ 0\ldots0\ldots,1\ldots1\ (k+\ell\text{ times})\ 0\ldots0\ldots). \tag{5}$$

Applying bit-invariance with $S=\{1,\ldots,k\}$, we have

$$d(1\ldots1\ (k\text{ times})\ 0\ldots0\ldots,1\ldots1\ (k+\ell\text{ times})\ 0\ldots0\ldots)=$$
$$d(0\ldots0\ (k\text{ times})\ 0\ldots0\ldots,0\ldots0\ (k\text{ times})\ 1\ldots1\ (\ell\text{ times})\ 0\ldots0\ldots). \tag{6}$$

By applying a permutation π that places numbers from $k+1$ to $k+\ell$ in the first ℓ positions, we conclude that

$$d(1\ldots1\ (k\text{ times})\ 0\ldots0\ldots,1\ldots1\ (k+\ell\text{ times})\ 0\ldots0\ldots)=$$
$$d(0\ldots0\ldots,1\ldots1\ (\ell\text{ times})\ 0\ldots0\ldots). \tag{7}$$

So, this term is equal to $d_\ell^{\neq}$, and the inequality (5) takes the desired form

$$d_{k+\ell}^{\neq}\le d_k^{\neq}+d_\ell^{\neq}.$$

5.2°. Let us first prove the second of the two inequalities (3).

Triangle inequality implies that

$$d(00\ldots0\ldots,0\ldots0\ (k-\ell\text{ times})\ 1\ldots1\ldots)\le$$
$$d(00\ldots0\ldots,0\ldots0\ (k\text{ times})\ 1\ldots1\ldots)+$$
$$d(0\ldots0\ (k\text{ times})\ 1\ldots1\ldots,0\ldots0\ (k-\ell\text{ times})\ 1\ldots1\ldots). \tag{8}$$

The first two distances in this equality are $d_{k-\ell}^{=}$ and $d_k^{=}$, so

$$d_{k-\ell}^{=}\le d_k^{=}+d(0\ldots0\ (k\text{ times})\ 1\ldots1\ldots,0\ldots0\ (k-\ell\text{ times})\ 1\ldots1\ldots). \tag{9}$$

Applying bit-invariance with $S=\{k-\ell+1,k-\ell+2,\ldots\}$, we have

$$d(0\ldots0\ (k \text{ times})\ 1\ldots1\ldots, 0\ldots0\ (k-\ell \text{ times})\ 1\ldots1\ldots) =$$

$$d(0\ldots0\ (k-\ell \text{ times})\ 1\ldots1\ (\ell \text{ times})\ 0\ldots0\ldots, 0\ldots0\ldots). \qquad (10)$$

By applying a permutation π that places numbers from $k-\ell+1$ to k in the first ℓ positions, we conclude that

$$d(0\ldots0\ (k-\ell \text{ times})0\ldots0\ (\ell \text{ times})\ 0\ldots0\ldots, 0\ldots0\ (k-\ell \text{ times})\ 0\ldots0\ldots) =$$

$$d(1\ldots1\ (\ell \text{ times})\ 0\ldots0\ldots), 0\ldots0\ldots). \qquad (11)$$

So, this term is equal to $d^{\neq}_{\ell}$, and the inequality (9) takes the desired form

$$d^{=}_{k-\ell} \le d^{=}_{k} + d^{\neq}_{\ell}.$$

The proposition is thus proven.

3 Auxiliary Result: Why Do We Need Metrics Which Are Not Permutation-invariant?

So why use existing metrics? Since it is possible to have permutation-invariant metrics, why do we need metrics like (1) and (2) which are not permutation-invariant? It turns out that this need comes from yet another natural requirement: computability.

Additional requirement: computability. At any given moment of time, we have finitely many observations and measurements, i.e., we only know finitely many bits of the sequences x and y. Thus, we need to estimate the distance based on this information. A natural requirement is that, as we get more and more bits, we should get a closer and closer approximation to the actual distance $d(x, y)$ – in other words, that for any desired accuracy ε, we should be able to find a natural number n so that, by knowing n first bits of each of the two sequences x and y, we should be able to estimate the distance $d(x, y)$ with accuracy ε.

Let us formulate this requirement in precise terms.

Definition 7. *We say that a metric $d : B \times B \to \mathbb{R}^+_0$ is* computable *if for every positive real number $\varepsilon > 0$ there exists an integer n such that if $x_1 \ldots x_n = X_1 \ldots X_n$ and $y_1 \ldots y_n = Y_1 \ldots Y_n$, then $|d(x, y) - d(X, Y)| \le \varepsilon$.*

Discussion. One can check that the usual – non-permutation-invariant metric (1) is computable: we can take n for which $2^{-n} \le \varepsilon$, i.e., $n \ge -\log_2(\varepsilon)$. Similarly, each metric (2) is also computable.

Let us prove, however, that a non-trivial permutation-invariant metric cannot be computable.

Definition 8. *We say that a metric is* trivial *if there exists a constant* $d_0 > 0$ *such that* $d(x, x) = 0$ *and* $d(x, y) = d_0$ *for all* $x \neq y$.

Proposition 2. *The only computable monotonic bit-invariant permutation-invariant metric on the set of all infinite binary sequence is the trivial metric.*

Proof. Let $d(x, y)$ be a computable monotonic bit-invariant permutation-invariant metric on the set of all infinite binary sequences. THis metric has the form described in the formulation of Proposition 1.

Let x and y be any two different sequences. Let $\varepsilon > 0$ be any positive real number. By definition of continuity, there exists an integer n for which if $x_1 \dots x_n = X_1 \dots X_n$ and $y_1 \dots y_n = Y_1 \dots Y_n$, then $|d(x, y) - d(X, Y)| \le \varepsilon$. Let us take:

- as X, an infinite sequence that continues the starting fragment $x_1 \dots x_n$ of the sequence x with all zeros, i.e.,

$$X = x_1 \dots x_n 00 \dots 0 \dots,$$

  and
- as Y, an infinite sequence that continues the starting fragment $y_1 \dots y_n$ of the sequence y with an infinite repetition of 01's, i.e.,

$$Y = y_1 \dots y_n 0101 \dots 01 \dots$$

Then, X and Y have infinitely many common bits and infinitely many different bits, so $d(X, Y) = d_\infty$. Thus, we have $|d(x, y) - d_\infty| \le \varepsilon$.

This is true for every ε, so we have $d(x, y) = d_\infty$. Thus, the metric $d(x, y)$ is indeed trivial, with $d_0 = d_\infty$. The proposition is proven.

Possible future work. In principle, instead of considering all possible permutations – as we did – we can consider only computable measure-preserving permutations; see, e.g., [3]. It would be interesting to analyze which metrics are invariant with respect to such permutations.

Acknowledgments

This work was supported in part by the National Science Foundation grants 1623190 (A Model of Change for Preparing a New Generation for Professional Practice in Computer Science), and HRD-1834620 and HRD-2034030 (CAHSI Includes), and by the AT&T Fellowship in Information Technology.

It was also supported by the program of the development of the Scientific-Educational Mathematical Center of Volga Federal District No. 075-02-2020-1478, and by a grant from the Hungarian National Research, Development and Innovation Office (NRDI).

The work of Irina Perfilieva was supported by European Regional Development Fund (ERDF) and European Social Fund (ESF) via the project "Centre for the development of Artificial Intelligence Methods for the Automotive Industry of the region" No. CZ.02.1.01/0.0/0.0/17_049/0008414.

The authors are thankful to all the participants of the 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'2022, (Milan, Italy, July 11–15, 2022), especially to Alexander Šostak, for valuable discussions.

References

1. R. Bēts, A. Šostak, and E. M. Miķelsons, "Parameterized metrics and their applications In word combinatorics", *Proceedings of the 19th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems IPMU'2022*, Milan, Italy, July 11–15, 2022, Vol. 1, pp. 270–281.
2. C. S. Calude, H. Jürgensen, and L. Staiger, "Topology on words", *Theoretical Computer Science*, 2009, Vol. 410, No. 24–25, pp. 2323–2335.
3. M. Li and P. Vitanyi, *An Introduction to Kolmogorov Complexity and Its Applications*, Springer, Berlin, Heidelberg, New York, 2008.